



Blind forecasting for Gaussian time-series

Thibault Espinasse, Fabrice Gamboa, Jean-Michel Loubes

► To cite this version:

Thibault Espinasse, Fabrice Gamboa, Jean-Michel Loubes. Blind forecasting for Gaussian time-series. 42èmes Journées de Statistique, 2010, Marseille, France, France. inria-00494759

HAL Id: inria-00494759

<https://inria.hal.science/inria-00494759>

Submitted on 24 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PRÉDICTION AVEUGLE DE SÉRIES CHRONOLOGIQUES GAUSSIENNES

Thibault Espinasse & Fabrice Gamboa & Jean-Michel Loubes

118 route de Narbonne 31062 Toulouse Cedex 9

Mots clefs : Statistique mathématique

1 Résumé :

Dans de nombreuses situations, le statisticien observe un échantillon fini issu d'un phénomène temporel X_1, \dots, X_n . Une modélisation usuelle est de considérer cet échantillon comme issu d'un processus stationnaire $\mathbf{X} := (X_t)_{t \in \mathbb{Z}}$ au second ordre (voir [2], [10]). Cela signifie que la variable aléatoire X_t est, pour tout $t \in \mathbb{Z}$, de carré sommable, et que la moyenne (que l'on suppose égale à 0) et la structure de covariance sont invariantes par translation. La covariance $\mathbb{E}(X_t X_s)$ dépend donc seulement de la distance entre t et s .

Le cas Gaussien est particulièrement simple, car le processus est alors aussi fortement stationnaire, ce qui signifie que

$$(X_1, \dots, X_n) \stackrel{\mathcal{L}}{=} (X_{t+1}, \dots, X_{t+n}), \quad (t \in \mathbb{Z}, n \in \mathbb{N}).$$

En particulier, tous les moments généralisés de ce processus multidimensionnel dépendent seulement des deux premiers moments. Le cas des processus Gaussiens est particulièrement adapté à la prédiction, car la régression linéaire fournit le meilleur prédicteur.

Il est donc aisé de construire, connaissant la structure de covariance du processus, l'opérateur de projection des valeurs à prédire (par exemple X_0) sur les valeurs observées (par exemple $X_{-N}, \dots, X_{-1}, N > 0$). De plus, sous certaines hypothèses de régularité, ce projecteur converge vers l'opérateur de projection $\text{Proj}_{H_{\mathbb{Z}^-}^X}$ sur le passé infini $(X_i)_{i < 0}$.

Dans cet exposé, on cherche à résoudre le problème de prédiction aveugle, c'est à dire lorsque la covariance est inconnue et que l'on doit, avec un unique échantillon, fournir à la fois un estimateur de la covariance, et un prédicteur des données manquantes. Ces deux problématiques de prédiction et d'estimation sont classiques (voir par exemple [7], [1], [3]), mais le cas aveugle l'est beaucoup moins. Bickel a prouvé la convergence d'un estimateur dans le cas d'observations indépendantes de la série chronologique [4]. En réalité, cet estimateur est performant dans le cas d'un unique échantillon. Notre méthode utilise cet estimateur pour construire, en aveugle, et avec un unique échantillon, un estimateur de l'opérateur de projection sur le passé infini.

La vitesse de convergence de cet estimateur sera cherchée, non pas en norme opérateur, mais en norme quadratique pour l'image (il s'agit en réalité du biais de l'erreur de

prédiction). Cela pose des problèmes supplémentaire de dépendance, car en réalité il s'agit de fournir une prédiction de variables manquantes avec l'échantillon même qui a été utilisé pour l'estimation.

Pour faire cela, utilisant une forme régularisée de l'estimateur de la covariance ainsi qu'une astucieuse décomposition de Schur (qui donne une expression équivalente à celle de Bondon obtenue dans [5] et [6]), nécessaire au calcul du biais, nous fournissons un estimateur de l'opérateur de projection sur le passé infini. La vitesse de convergence est obtenue en fonction de la régularité de la densité spectrale du processus, à partir d'un résultat de concentration du à Comte [8].

2 Abstract:

In many concrete situations the statistician observes a finite path X_1, \dots, X_n of a real temporal phenomena. A common modeling is to assume that the observation is a finite path of a second order weak stationary process $\mathbf{X} := (X_t)_{t \in \mathbb{Z}}$. This means that the random variable (r.v.) X_t is, for any $t \in \mathbb{Z}$, square integrable and that the mean (supposed to be equal to zero) and the covariance structure of the process is invariant by any translation on the time index. That is, for any $t, s \in \mathbb{Z}$, $\mathbb{E}(X_t) = 0$ and $\mathbb{E}(X_t X_s)$ only depends on the distance between t and s . A more popular frame is the Gaussian case where the additional Gaussianity assumption on all finite marginal distributions of the process $(X_t)_{t \in \mathbb{Z}}$ is added. In this case, as the multidimensional Gaussian distribution only depends on moments of order one and two, the process is also strongly stationary. This means that the law of all finite dimensional marginal distributions are invariant if the time is shifted:

$$(X_1, \dots, X_n) \stackrel{\mathcal{L}}{=} (X_{t+1}, \dots, X_{t+n}), \quad (t \in \mathbb{Z}, n \in \mathbb{N}).$$

Gaussian stationary process are very popular because they share plenty of very nice properties concerning their statistical identification and prediction (see, for example, [2] or [10]). For instance, a well known property of Gaussian time series is that linear prediction is optimal. Hence, if one wish to predict for $t \geq 0$, X_t from X_{-N}, \dots, X_{-1} ($N > 0$) the r.v. minimising the prediction error is just a linear combination of X_{-N}, \dots, X_{-1} involving a linear projector operator onto the complete infinite past $\text{Proj}_{H_{\mathbb{Z}^-}^X}$ defined on the Hilbert space H_{∞}^X generated by the process \mathbf{X} . Furthermore, $\text{Proj}_{H_{\mathbb{Z}^-}^X}$ may be computed from the covariance function of the process. In this talk, we work with a Gaussian stationary process and we will address the problem of blind filtering. This means that observing the finite sample path X_{-N}, \dots, X_{-1} of the process we wish to predict the future values X_t , $t \geq 0$ without knowing the covariance structure (blind means that the covariance of \mathbf{X} is unknown). The prediction problem is classical (see for example [7],[1], [3]). The blind one is less classical and have been very few studied. The particular case of Kriging (see [9], [10]) in which a parametric model on the covariance is assumed is discussed and studied in [10]. An interesting work of Bickel and al [4] consider the case where many

samples are available. In our work, we do not assume any parametric model and set and solve the problem in a nonparametric way. Our method relies on the estimation of the covariance function. This function is estimated and an empirical regularisation method jointly with a clever Schur decomposition of the inverse covariance operator allow to build an accurate estimate of $\text{Proj}_{H_{\mathbb{Z}^-}^X}$. The asymptotic property of this estimate is study and rate of convergence are stated. Roughly speaking, the rate of convergence stated in the main Theorem depends of the regularity of the covariance function through the spectral density of the process (one of the main assumption is that the spectral measure of \mathbf{X} has a density).

References

- [1] Anestis Antoniadis, Efstathios Paparoditis, and Theofanis Sapatinas. A functional wavelet-kernel approach for time series prediction. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 68(5):837–857, 2006.
- [2] Robert Azencott and Didier Dacunha-Castelle. *Series of irregular observations*. Applied Probability. A Series of the Applied Probability Trust. Springer-Verlag, New York, 1986. Forecasting and model building.
- [3] R. J. Bhansali. Parameter estimation and model selection for multistep prediction of a time series: a review. In *Asymptotics, nonparametrics, and time series*, volume 158 of *Statist. Textbooks Monogr.*, pages 201–225. Dekker, New York, 1999.
- [4] Peter J. Bickel and Elizaveta Levina. Regularized estimation of large covariance matrices. *Ann. Statist.*, 36(1):199–227, 2008.
- [5] Pascal Bondon. Prediction with incomplete past of a stationary process. *Stochastic Process. Appl.*, 98(1):67–76, 2002.
- [6] Pascal Bondon. Influence of missing values on the prediction of a stationary time series. *J. Time Ser. Anal.*, 26(4):519–525, 2005.
- [7] Sofiane Brahim-Belhouari and Amine Bermak. Gaussian process for nonstationary time series prediction. *Comput. Statist. Data Anal.*, 47(4):705–712, 2004.
- [8] Fabienne Comte. Adaptive estimation of the spectrum of a stationary Gaussian sequence. *Bernoulli*, 7(2):267–298, 2001.
- [9] Arthur S. Goldberger. Best linear unbiased prediction in the generalized linear regression model. *J. Amer. Statist. Assoc.*, 57:369–375, 1962.

- [10] Michael L. Stein. *Interpolation of spatial data*. Springer Series in Statistics. Springer-Verlag, New York, 1999. Some theory for Kriging.